Technical Report CADD-98-XX
December 1998

Tri-Service CADD/GIS Technology Center

Tri-Service Guidelines for Electronic Document Management Systems (EDMS) for Facility Management

Approved for Public Release; Distribution is Unlimited

Prepared for Tri-Service CADD/GIS Technology Center
U.S. Army Corps of Engineers, Waterways Experiment Station
Information Technology Laboratory

Technical Report CADD-98-XX
December 1998

Tri-Service Guidelines for Electronic Document Management Systems (EDMS) for Facility Management

by Robert J. Hanson, John A. Owens, James R. Daley
Michael Baker Jr., Inc.

Andrew J. Synnott
TSA/ADVET

Laurel T. Gorman
Tri-Service CADD/GIS Technology Center

Prepared for Tri-Service CADD/GIS Technology Center
U.S. Army Corps of Engineers, Waterways Experiment Station,
Information Technology Laboratory

Under Delivery Order No. 8 Under Contract No. DACW3996-D-0017

Contents

List of Tables

List of Figures

Preface

This report provides an overview of Electronic Document Management Systems for engineering and facility management documents and conveys lessons learned by organizations using EDMS for these documents. Guidelines for evaluating EDMS implementation, EDMS requirements, and commercially available hardware and software also are presented.

This report is a product of the Tri-Service CADD/GIS Technology Center 1997 Project Number 62, Engineering Document Management System. The project was funded and conducted by the Tri-Service CADD/GIS Technology Center, Information Technology Laboratory (ITL), U.S. Army Engineer Waterways Experiment Station (WES). The Tri-Service CADD/GIS Technology Center was chartered in 1992 to promote the use of CADD and GIS technologies for life-cycle facilities management within the Army, U.S. Army Corps of Engineers, Navy, and Air Force. The Center operates under the guidance of Dr. N. Radhakrishnan, Director, ITL, and Mr. Harold Smith, Chief, Tri-Service CADD/GIS Technology Center. The Center functions under the guidance of several oversight committees including the Executive Working Group (EWG), Field Technology Advisory Group (FTAG) and Facility Management Working Group (FMWG). Members of these groups are listed below.

Executive Working Group Membership		
Name	**Membership**	**Affiliation**
M. K. Miles	Chair	Corps of Engineers
Mikeual Perritt	Member	Air Force
Peter J. Sabo	Member	Army
Fredrik (Rik) Wiant	Member	Army
Ron Hatwell	Member	Corps of Engineers
N. Radhakrishnan	Member	Corps of Engineers
Thomas M. Karst	Member	Defense Logistics Agency
Jim Carberry	Member	Navy
Nancy Blyler	Alternate	Corps of Engineers
Thomas R. Rutherford	Member	OSD
Jim Whittaker	Member	OSD
Paul Herold	Member	Coast Guard
Bobby Bean	Member	Navy
Dana (Deke) Smith	Member	Navy

All EDMS projects are performed under the direction of the Facility Management Field Working Group (Facility Management FWG). This FWG represents resources from two previous FWGs: Maintenance and Repair, and Space Utilization. The two charts that follow list names of members of the Field Technology Advisory Group and Facility Management FWG.

Field Technology Advisory Group (FTAG) Membership		
Name	**Membership**	**Affiliation**
Bobby Bean	Chair	Navy
Randy Lierly	Member	Air Force
Victoria Williams	Member	Air Force
Jim Butler	Member	Army
Deborah Duncan	Member	Army
Phil O'Dell	Member	Corps of Engineers
Eugene Tickner	Member	Corps of Engineers
Thomas M. Karst	Member	Defense Logistics Agency
Carolyn Wilber	Member	Navy
Robert Wood	Member	Navy

Facility Management Field Working Group		
Name	**Membership**	**Affiliation**
Laurel Gorman	Facilitator	Tri-Service Center
Nancy Towne	Facilitator	Tri-Service Center
Betty Marchbanks	Member	Air Force
Marta Reiner	Vice Chair	Air Force
David Carr	Member	Army
Bob Riley	Member	Army
Ray Consoli	Chair	Corps of Engineers
Jeff Bryant	Member	Navy
Bill Hudson	Member	Navy
Vivian Sanchez	Member	Navy

To establish procedures and priorities for EDMS issues yet to be addressed under the overall EDMS initiative, a Task Group was formed. The Task Group is led by Ms. Vivian Sanchez, NAVFACENGCOM, Southwestern Division, who is a member of the Facilities Management FWG. Other members include Ms. Marta Reiner, USAF, member of FM FWG; Mr. Roger Porzig, COE, FY98/99 Chair of Systems Group, and Mr. Gary Boyd, NAVFACENGCOM, Southern Division, member of the Design FWG. Proponents who provided guidance and detailed reviews of the draft documents were Ms. Jean McGinn, from HQ USACE and Mr. Bobby Bean, Director of Public Works, from Patuxent River Naval Air Station.

The Corps of Engineers Principal Investigator and Field Working Group (FWG) Facilitator was Ms. Laurel T. Gorman, P.G.

Michael Baker Jr., Inc. served as the prime consultant for this Task Order. TSA/ADVET served as a subconsultant to Michael Baker Jr., Inc. on this project. Key Michael Baker Jr., Inc. personnel involved in production of this report were Mr. Robert J. Hanson, Project Manager; Mr. James R. Daley, Principal Investigator, and Mr. John A. Owens, Systems Analyst. The primary TSA/ADVET contributor was Mr. Andrew J. Synnott.

Members of the Executive Working Group, Facility Management Field Working Group, Field Technology Advisory Group and Task Group contributed to this effort. Many other individuals, representing numerous organizations, also made substantial contributions to the project. The Tri-Service CADD/GIS Technology Center appreciates and acknowledges those persons and organizations listed on the following table that contributed to the report effort. Many contributed first-hand knowledge and experiences which do not yet exist in the public domain. Others offered critiques of the interim deliverables. All added value to the report.

Person	Organization	Primary Contribution(s)
Mr. Bobby Bean	Naval Air Station Patuxent River Public Works Department	Met to discuss the PAX EDM, got supplier (NTA) to participate, supplied presentation materials, edited PAX information in report, and reviewed some interim deliverables
Mr. Larry Condry	Baltimore Gas & Electric Gas Maps & Records Unit	Completed a user survey and furnished additional information via personal communication
Mr. Ray Consoli	U.S. Army Corps of Engineers Center for Public Works	Reviewed and commented on interim deliverables.
Mr. Dan Jave	Iowa Ntl Guard, Johnston IA Facilities & Construction Office	Completed a user survey
Mr. Edwin W. Kincaid	Tinker Air Force Base, OK Air Logistics Command	Personal communication and presentation materials on Tinker's PDM
Mr. Wayne Kuenzli	Naval Supply Systems Command, Mechanicsburg, PA	Personal communication with JEDMICS PMO regarding current status of the system
Mr. Reed MacMillan	Aberdeen Proving Ground, MD Conservation and Restoration Div.	Completed survey and follow-up site visit regarding Aberdeen's ERPMS
Ms. Connie H. Malarkey	DOE Oak Ridge Centers for Manufacturing Technology	Personal communication regarding DOE's EDMS
Ms. Louise McMonegal	HQ, Naval Facilities Engineering Command	Supplied information on NAVFAC ongoing EDMS guidelines effort
Mr. James Michonski	Norfolk NSY, Portsmouth, VA Public Works Center	Completed a user survey
Mr. Roger W. Porzig	U.S. Army Corps of Engineers Jacksonville District	Completed user survey and participated in on-site follow-up meeting. Also furnished EDMS presentation slides
Mr. Ralph Scheid	U.S. Army Corps of Engineers New Orleans District	Reviewed 90% deliverable and furnished comments
Ms. Debbie Tankersley	The Analytical Sciences Corp. (TASC, Inc.)	Personal communication regarding EDMS use for procurements at Warner Robins AFB

Conversion Factors Table

Conversion Factors

Eight Bits	=	1 Byte
1024 Bytes	=	1 Kilobyte
1024 Kilobytes	=	1 Megabyte
1024 Megabytes	=	1 Gigabyte
1024 Gigabytes	=	1 Terabyte
1024 Terabytes	=	1 Petabyte
1024 Petabytes	=	1 Exabyte

Decimal/Binary Comparison

Prefix	Decimal	Binary
kilo-	1000^1	1024^1 = 2^10 = 1,024
mega-	1000^2	1024^2 = 2^20 = 1,048,576
giga-	1000^3	1024^3 = 2^30 = 1,073,741,824
tera-	1000^4	1024^4 = 2^40 = 1,099,511,627,776
peta-	1000^5	1024^5 = 2^50 = 1,125,899,906,842,624
exa-	1000^6	1024^6 = 2^60 = 1,152,921,504,606,846,976

Definition of Terms

Note: Many of these definitions were derived from the "Pocket Glossary of Computer Terms and Definitions" (Black Box Corporation, 1997) and "PC Webopedia Encyclopedia" (PC Webopedia, Undated).

Bandwidth	The transmission capacity of a communications line
Bit	The smallest element of computer storage
BMP	Windows and OS/2 raster graphics
Byte	A unit of computer storage composed of eight bits
CAD	Computer-Aided Design
CADD	Computer-Aided Drafting and Design
CAE	Computer-Aided Engineering
CGI	Computer Graphics Interface graphics language for screens, printers, and plotters
CCITT	Comité Consultatif International Téléphonique et Télégraphique, an organization that sets international communications standards (now known as ITU)
Client	A workstation or PC in a client/server environment
COLD	COLD originally stood for Computer Output to Laser Disk. Today, it more appropriately stands for Computer On-Line Data. Either way, it is a technology that provides for the storage and retrieval of computer generated reports and documents. COLD documents are generated from a user's own computer system and captured as data.
DGN	Intergraph or Bentley CADD file format
Digitize	The process of converting an image or signal into digital code
Dot Pitch	The distance between dots on a color monitor
DPI	Dots Per Inch - the measurement of printer resolution
DWG	AutoCAD CADD file format
EDMS	Electronic Document Management System
EDMSs	Electronic Document Management Systems
Ethernet	A local area network (LAN) developed by Xerox
Firewall	A network node limiting segment traffic. Also implemented for security purposes.
FTP	File Transfer Protocol - in a TCP/IP network a set of commands used to log onto a network, list directories, and copy files
GIF	Graphics Interface Format - Compuserve raster graphics
GB	Gigabyte - 1,073,741,824 bytes or 1,024 megabytes

GUI	Graphic User Interface - a program interface that takes advantage of the computer's graphics capabilities to make the program easier to use
IP	Internet Protocol - the protocol used in gateways to connect networks at the OSI (Open System Interconnection) network level
JPEG	Joint Photographic Experts Group: an ISO/ITU standard for compressing still images
Jukebox	System containing many separate individual storage units (such as CD-ROMs or optical disks) that can be accessed one at a time
KB	Kilobyte – 1,024 bytes
LAN	Local Area Network - a communication network composed of servers, workstations, and network operating system limited to a specific geographical area.
MB	Megabyte - 1,048,576 bytes or 1,024 kilobytes
Metadata	Descriptive information pertinent to a document – i.e., date, version, title, etc.
PCX	Zsoft Corporation raster graphics
PDF	Portable Document Format, a file format developed by Adobe Systems
Pixel	The smallest element or dot on a video display
PostScript	An Adobe text/graphics programming language
RAID	Redundant Array of Inexpensive Disks - a method of storing data on multiple hard-drives
Raster	A rectangular pattern of scanning lines that produce a digital image
Scanner	A device that creates digital images of documents
SCSI	Small Computer Standard Interface - a hardware interface that supports from 7 to 15 peripheral devices
Server	A computer or device on a network that manages network resources
TCP/IP	Transmission Control Protocol/Internet Protocol - a set of layered protocols that enables shared applications among PC's, hosts, or workstations in a high-speed communications environment
Terabyte	1,099,511,627,776 bytes or 1,024 gigabytes
TIFF	Tagged Image File Format - raster graphics format developed by Aldus and Microsoft
Vector	Computer graphics represented by points, lines, and other geometric entities
WAN	Wide Area Network - a computer network that spans a relatively large geographical area. Typically, a WAN consists of two or more local-area networks (LANs).
Workflow	The automatic routing of documents to the users responsible for working on them

Workstation	A high-performance single-user microcomputer or minicomputer that is used for graphics, CAD, CAE, simulation, and scientific applications.
WORM	Write Once Read Many - An optical disk that can be recorded only once.
WYSIWYG	What You See Is What You Get. A WYSIWYG is a word processor that enables the user to see on the display screen exactly what the text will look like when printed.

Executive Summary

The focus of this report is on facility management applications for Electronic Document Management Systems (EDMSs). Because of the significant differences in managing large-scale, multiple-level, often native CADD files used by facility managers, as opposed to small-scale business documents, some software vendors have adopted another meaning for their EDM systems - - Engineering Document Management Systems. The differences in business versus engineering document EDMSs are document size, and content, and primary functions. Many EDMSs are implemented as means to scan static (released state) hard-copy documents into an electronic repository, from which they can be accessed using the improved file location and distribution capabilities offered by an effective EDMS. In these applications, for instance, creation of an electronic technical library, the usefulness of an EDMS is judged largely on its ability to effectively capture, store, and distribute information. However, an EDMS capable of supporting an engineering design or facility management role also must have a strong workflow capability and fully support file manipulation via native applications. The workflow component must be especially strong when the EDMS is used with engineering documents that are in-process as opposed to released.

Research shows that the Department of Defense has needs that mirror those in the private sector. For example, the DoD has adopted a proprietary EDMS known as the Joint Engineering Data Management Information and Control System (JEDMICS). This system has robust image capture, storage, and distribution capabilities because it was created to serve as a "DOD repository environment with 47 Army, Air Force, Marine Corps, DLA and Navy sites" (Redstone Arsenal, 1997a and b). Indeed, with a reported 55 million images transitioned to JEDMICS, a huge repository has been created. Although smaller in size, many public utility and private sector clients have used EDMSs to create readily accessible document repositories. For instance, one of the survey respondents, Baltimore Gas and Electric Company, scans current utility drawings to capture, store, and distribute up-to-date information to its district offices. In many cases, an EDMS must give end-users the ability to view and/or print current documents without a capability to alter those documents in a native application. The ability to electronically store and retrieve engineering documents offers numerous benefits to private sector firms. As an example, U.S. Steel uses an EDMS at its Edgar Thompson Works to provide direct, electronic access to its engineering drawing database from the plant floor (Formtek, undated). In U.S. Steel's case, quick access and reduced paper shuffling are decreasing equipment downtime, improving profitability.

Use of EDMSs in engineering and facility management applications is complex, largely due to the

- Much larger sized files than standard business documents
- Linkage of CADD files to reference or seed files, per tables, color tables and font reference files
- Iterative review and drawing/discipline coordination process that need a strong workflow process
- Variability in workflow, dictated by the content of individual drawings and the persons contributing to, or reviewing and redlining the drawings
- Extended life-cycle of documents when carried through from the engineering stages to the bidding, construction, as-built, facility management and subsequent rehabilitation and eventual demolition stages.

Given these disparities in the EDMS requirements for business documents versus engineering documents, organizations contemplating acquisition of an EDMS must perform a thorough analysis of its needs and expectations. Organizations must develop a vision of how EDMS will be used as a productivity improvement tool. Enterprise-wide EDMSs are a concept which may be achievable with commercially available hardware and software. However, achieving such a goal may require acquisition of two different EDMSs, one suited to the capture, storage, and distribution of legacy, often hard-copy, business documents and the other suited to engineering and facility management applications. In any case, a phased approach to implementation is needed.

This report presents a number of other specific recommendations for implementing an EDMS that meets an organization's expectations. Eight of the more important recommendations are listed below.

- Clearly understand existing systems and workflows to ensure that candidate EDM systems can be integrated into your organization.
- Establish a clear purpose and objectives for your EDMS. Identify the types of documents the system must handle, how files will be entered into the EDMS, and how and if legacy data and systems are to be addressed.
- Recognize that cultural acceptance issues will arise and foster acceptance by involving stakeholders early and often. Some persons resist change and some may even subvert process enhancements.
- Take a calculated, methodical approach to EDMS implementation. Solve simple, bounded, and measurable processes first, ensuring that these processes are representative of the complexity of other processes within the organization. Select those processes that allow for an easy and measurable return on investment.

- Use a phased approach by implementing a pilot project to prototype the proposed system. This approach will ensure success and help to define the metrics used to demonstrate success.

- Incorporate incremental versus revolutionary changes in workflow. Choose an EDMS that can digitally enable or automate existing manual processes rather than require existing processes to be re-engineered.

- Remember that training will be needed during both pilot and production phase implementation and periodically thereafter. Computer-based training can be very attractive.

- Be certain that a selected EDMS supplier fully understands your needs and expectations and has satisfactorily served other customers with similar needs and expectations in the past.

DoD may elect to conduct one or more well-documented demonstration projects to show users how to identify their specific needs, identify and select a vendor, establish an acceptance plan, and implement the EDMS following a phased approach, including training at appropriate stages. To the extent possible, all tangible benefits of implementing an EDMS should be quantified and compared to EDMS costs to determine the return-on-investment (ROI).

1 Introduction

Background

Electronic Document Management Systems (EDMSs) consist of the computer hardware and software that allow for the integrated preparation, input, distribution, storage, location, and retrieval of electronic documents, whether created initially electronically or converted from paper documents.

Continued advancements in computer hardware and software, including the proliferation of client/server networks, have made EDMSs possible. Although technological advances enabled development of EDMSs, demand for these systems has been driven by the need to intelligently manage an ever-growing warehouse of documents and data. Information is readily available on EDMSs that can manage traditional (non-graphic data, such as text files) business documents. Large-scale, graphic documents important to facility managers present special challenges for EDMSs. Therefore, this report summarizes the status of EDMSs capable of managing multiple engineering and facility management document formats. It also provides DoD personnel with guidelines for planning and implementing an EDMS. However, this report intentionally does not compare COTS EDMS products because of the sheer number of suppliers, frequent supplier name changes from mergers and acquisitions in the computer hardware and software industry and nearly continuous product upgrades that would make a comparison obsolete within weeks or months.

Purpose and Need

The study purpose and need as contained in the Tri-Service CADD/GIS Technology Center's June 25, 1997, Statement of Work, are as follows:

> A critical need across DoD installations and USACE Civil Works and Military Programs projects is the ability to access engineering and facility management documents that are in a variety of formats including digital and hard-copy. It is difficult to locate, retrieve, store and distribute data contained in these documents with today's management and retrieval processes. Customers and users of this information traditionally copy required data into unique projects, creating large storage and duplicative update requirements, and loss of an accountable edit trail. As electronic documents are being added to the inventory of engineering documents at DoD offices, the need to file and store these documents for easy management

and retrieval increases. Optimally, commercial EDMS solutions provide the framework to establish metadata (data about data) for each piece of data, enabling DoD to locate, track and access CADD drawings, geospatial data sets, computer-aided facility management information, and documents by accessing them. Additionally, EDMS solutions will save resources by allowing data to be entered only once and used many times. As installations implement EDMS technology, guidelines and resources will be critical to the adoption and sustainment of this technology.

Specific objectives for the investigation and the guidelines are:

- Evaluate EDMSs currently in use, or being installed, at various DoD offices
- Document lessons learned in using EDMS technologies
- Develop criteria and guidelines for determining EDMS requirements and evaluating commercial EDMS hardware and software

What are EDMSs?

Definition

The term Document Management is used in so many ways that confusion reigns when trying to describe what it really means. Is it a technology that is used to manage the distributed repositories of documents now dispersed throughout many organizations? Is it the set of technologies that enable organizations to disseminate information to its internal resources, its clients, and its suppliers? Is it the set of technologies such as imaging and forms processing that allow organizations to input and retrieve these paper-based documents in a convenient way? Is it technologies like workflow and groupware that manage both the transaction-oriented and collaborative ways that documents should be processed within an organization? Is it the non-technical management issues that organizations need to address to effectively process their organizational memory?

The following paragraphs give the definitions used by several leading organizations in the EDMS industry.

AIIM International (AIIM International, undated) defines document managers as the products and services that provide revision control and repository-oriented services for the electronic documents located throughout an organization. Nevertheless, effective document management includes the use of these types of products and more. The integration of imaging, workflow, groupware, document managers, optical character recognition and other technologies, together with realistic standards-compliance and intelligent organizational management of these documents are what make up effective document management.

The Gartner Group (The Gartner Group, 1993) defines document management as a highly integrated set of middleware services that integrate library services, document manufacturing, and document interchange with critical business process applications around a client/server topology using open application interfaces.

International Data Corporation (International Data Corporation, 1993) defines document management as a software system that is capable of organizing document production, managing accessibility and distribution of volumes of textual documents, and overseeing document flow.

Interleaf (Interleaf Inc., undated) defines document management as not only the technology that manages documents, but more important, that manages the information within documents. Interleaf regards document management as a set of software and services through which business-critical information is managed by enabling the creation, assembly, control and distribution of this information. Document management is about more than documents - it's about information and strategic business processes.

Given the diversity of these definitions, the project team arrived at this simple definition to focus the study effort: EDMSs consist of the computer hardware and software that allow for the integrated preparation, input, distribution, storage, location, and retrieval of electronic documents, whether initially created electronically or converted from paper documents. By considering the life-cycle of a typical engineering document, such as an engineering drawing, the required components of an effective EDMS become clear.

A well configured EDMS certainly improves enterprise-wide operating efficiencies. However, a careful analysis of requirements and expectations is needed to avoid disappointments in terms of:

- anticipated costs and returns
- cultural issues, including user acceptance
- technology limitations
- technology transition requirements

EDMS Features and Components

Engineering, planning, and facility management documents may be created manually (requiring a separate document scanning step) or electronically. In either case, an iterative process of internal reviews and approvals is typical. For example, “Final” engineering drawings emerge from this process. Individual drawings are bundled as needed for bidding and construction purposes. Equating this process to required EDMS features, an EDMS should to be able to store documents while in progress, allow manipulation via application software, distribute full documents to appropriate reviewers, and capture digital information from paper-based documents. Effective document management requires an ability to store and track document versions/revisions that result from the iterative review process, control file access, and store documents in a secure repository.

The ability to easily locate and retrieve documents is as important as the ability to store them, since stored information that cannot be easily or readily accessed can result in poorly executed design or construction projects or poorly managed facilities. Effective document retrieval requires linkage of metadata to the documents in a repository. Metadata enables EDMSs to locate documents for retrieval. Metadata is information about a document, such as the origination date, originator, document version, software version, content etc. Metadata can be thought of as database entries needed to retrieve a relevant document when only certain information, like the preparer, facility name or creation date is known. Depending upon design, the EDMS can require that pertinent metadata be entered before a file can be submitted to the repository.

Documents have an extended life-cycle beyond initial creation. Using the engineering drawing example, the life-cycle carries through to as-built drawings and baseline documents for future renovation, rehabilitation, or demolition efforts. This extended life cycle, characterized by evolution of documents over time, generates the need for an automated process that ensures changes in one document are reflected in related documents.

To gain a full understanding of EDMS, you can consider the individual components that comprise an EDMS and the primary functions that these various components perform. Because the individual pieces are building blocks of an EDMS, the following paragraphs introduce typical EDMS components. Recognizing that all components must interact to achieve efficient document management, Section 2 gives a detailed discussion of the typical EDMS components and primary functions an EDMS must perform or support.

Primary components and their associated function in an EDMS include:

- Input Devices These are the devices that directly, or through a conversion process, place documents into the EDMS. Any computer system can serve as an input device. Optical scanners are used to input non-digital or legacy hard-copy data to the EDMS. TIFF (tagged image file format); and Adobe's Acrobat PDF (portable document format) are the most commonly accepted image formats (Doculabs, 1997). Other accepted formats are BMP, GIF, JPEG.

- Storage Devices (Repository) The document repository "stores, controls and manages documents. Primary components include file or data servers, database servers, FTP servers, tape backups (9 track, 8mm, 4 mm, etc.), optical storage and CD ROM. These devices store information in a repository for future retrieval. Key repository functions include library services (e.g., controlling access to individual documents, document cataloging, check-in/check-out, and searching for and retrieving documents and version control, including a history of all instances of a document as it changes over time.)" (Boyle, 1997). Workflow systems are used to automate routing and processing functions, typically adhering to specifications developed by the Workflow Management Coalition (Doculabs, 1997).

- Retrieval and Distribution Devices Retrieval and distribution components consist of workstations, printers, plotters, recordable CD units, and facsimile machines. COLD (Computer Output to Laser Disk) technology is gaining popularity since it facilitates quicker and more secure document retrieval. Formatting and rendering packages are used to allow print-on-demand capabilities. Use of Internet and Intranet technologies make documents retrievable by a wide array of users, within security access limitations.

Available EDMS Guidance

The Tri-Service CADD/GIS Technology Center had the Jordani Consulting Group prepare an EDMS overview document, entitled "Electronic Document Management Systems" (Jordani Consulting Group, 1996). This and related reports can be downloaded from the following URL: http://tsc.wes.army.mil/projects/.

The Jordani report presents a 15-page high level overview of EDMS technology and applications. It contains:

- An overview and definition of EDMS
- Business challenges that an EDMS can help overcome
- EDMS features useful to any organization or business
- Criteria for selecting an EDMS
- Implementation issues
- Anticipated future advancement in EDMS functionality

As a primer, Jordani's report has appeal to a wide audience, especially those persons with limited prior knowledge of EDMS technologies.

Publications and World Wide Web addresses furnished in the references and bibliography sections also contain pertinent information on EDMSs in general, with a focus on engineering and facility management document applications for EDMS. The Tri-Service CADD/GIS Technology Center maintains numerous world-wide links to EDMS sites; these links can be accessed at http://tsc.wes.army.mil/links/.

2 EDMS Components

Introduction

The purpose of this section is to list and define the various components involved in the EDMS process. Additional technical specifications, requirements, and suggestions are included under the sub-heading "EDMS Components – Technical Specifications."

EDMS Components and Activities

Engineering document management encompasses many different activities within the life cycle or process workflow defined by the engineering process. The activities involved in EDMS can be placed into five categories: Capture, Store, Manipulate, and Distribute activities, linked by the Manage activity. Figure 1 illustrates typical EDMS architecture and the relationship of the five primary activities. The document workflow is delineated by the arrows surrounding the "Manage" or "EDMS" activity.

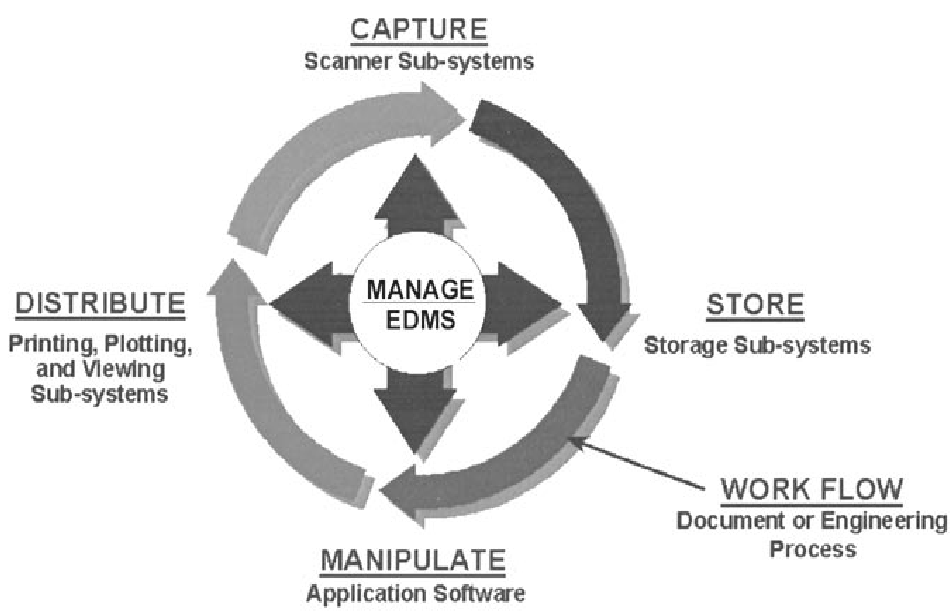

Figure 1. EDMS System Architecture

The EDMS activity must interact with each of the other four activities to allow for communication between all of the activities.

EDMS technology is supported by a network of hardware and software. The architecture of an EDMS consists of four primary elements: desktop clients, servers, databases, and system storage (Bielawski & Boyle, 1997).

Figure 2 shows a typical layout of components in the electronic engineering process. These components are defined in more detail later, specific to the engineering document management processes. The server platforms shown in Figure 2 are logical servers delineating server functions. Several or all of these server functions may reside on one physical hardware server device. Several physical server devices may also be used to manage the listed functions.

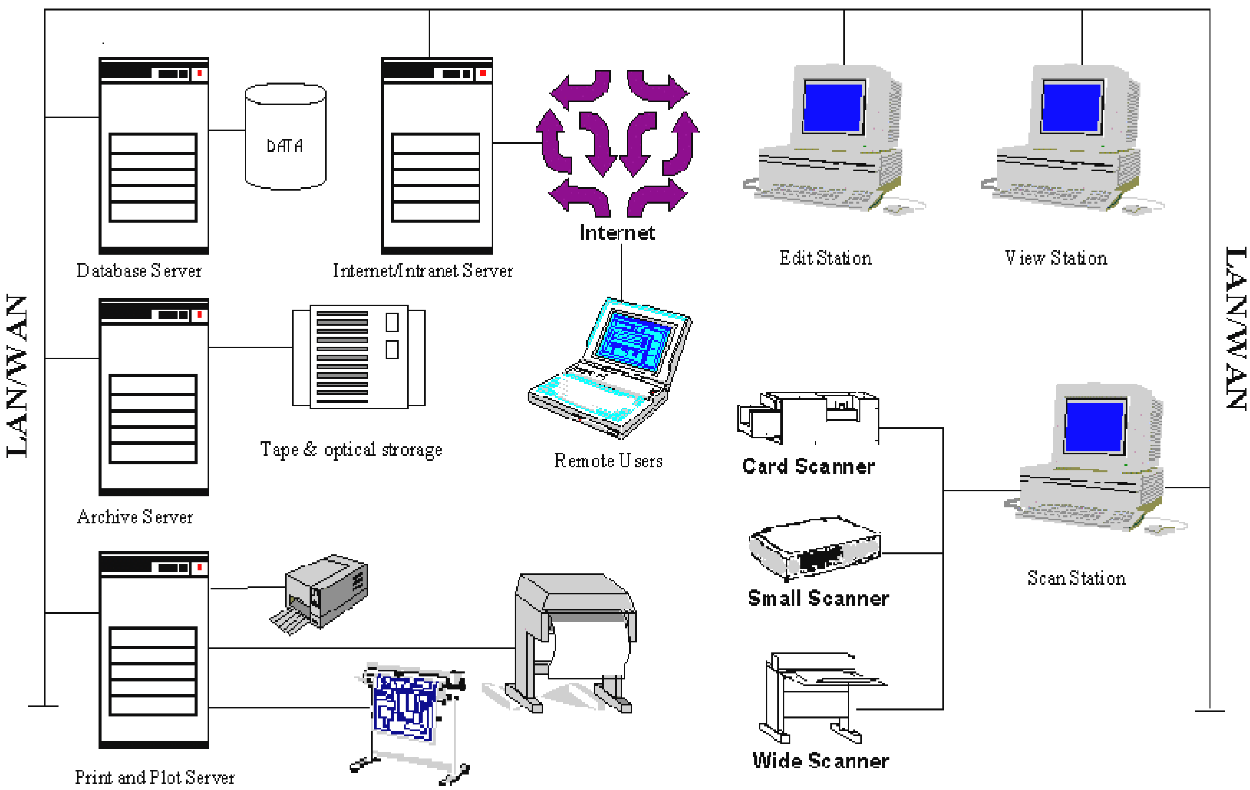

Figure 2. Components of the Electronic Engineering Process

Each of the different activities delineated in Figure 1 interact with one or more of the components delineated in Figure 2. Following are definitions of each activity and explanations of how the components interact with the EDMS.

Capture Activity

Introduction

According to the American Institute of Information Management, approximately 20% of the documents in use today are in digital form. Lately, significant efforts have been exerted to digitally capture the paper-based engineering document vaults. Many benefits accrue to users who expend the resources needed to digitally capture documents. In simplest terms, digitally capturing paper engineering documents is accomplished via scanning the paper documents (imaging) and indexing the scanned files into an EDMS.

An imaging subsystem consists of several key components (Bielawski & Boyle, 1997):

- Scanner
- Image Capture Software
- Image Processing Software
- Optical Character Recognition Software

Although some organizations adopt a "from this day forward" approach, using the EDMS to manage newly created documents, many organizations stress the ability to convert existing data warehouses into electronic repositories. Organizations that need to quickly convert hard copy information to electronic files need to carefully analyze requirements for the entire imaging subsystem.

In a paper-based engineering environment significant amounts of time are spent in simply managing the paper engineering data base rather than actually using the information stored on the paper. According to the Dephi Consulting Group (Delphi Consulting Group, 1997),

- 90% of all documents handled each day are merely shuffled
- Gathering and transferring paper documents consumes 90% of typical office tasks
- Workgroups lose 15% of all documents they handle
- 30% of the above workgroup's time is spent trying to find them

The above statistics indicate that handling paper-based information yields no "value added" time to the process of working with such information. Significant savings can be realized in improved personnel efficiencies. According to AIIM (American Institute of Information Management), companies spend 5% of their total filing costs on equipment, 20% on space, and 75% on salaries of those individuals spending time filing and requesting paper-based documents.

Figure 3 further graphically defines the traditional baseline situation using paper-based engineering document vaults. Key components of the imaging subsystem are described after Figure 3. The imaging subsystem allows organizations to move from manual vaults to electronic repositories.

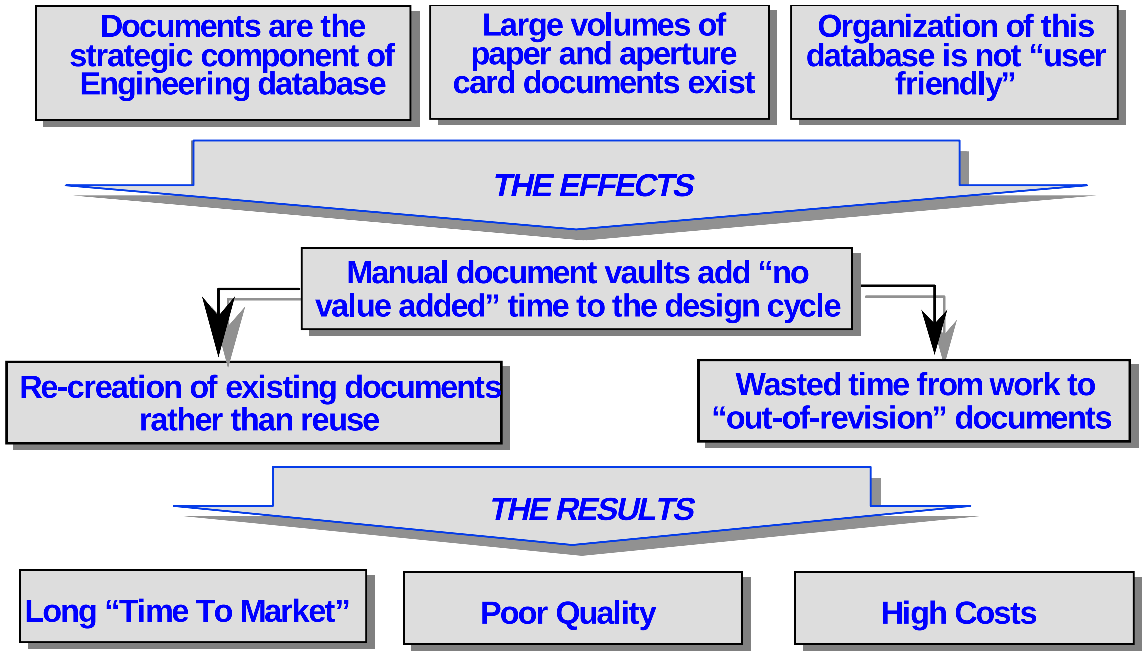

Figure 3. Paper-Based Engineering Document Vault - The Situation

Imaging Process

Scanner

Scanners are hardware devices that read hard-copy (legacy) documents. Scanners are available to generate black and white, gray-scale or color outputs with resolutions typically ranging from 200 to 1200 dots per inch. However, high resolution and color yield very large file sizes. Three scanner configurations are used, flat-bed, automatic sheet feed and roll-feed (used for large-scale engineering drawings). Aperature card scanners also are useful in an engineering environment.

Image Capture

Image capture software captures the scanned image and outputs in desired formats, most commonly TIFF. File compression also is a critical function of the image capture software due to the large file sizes associated with scanned documents.

Image Processing

The scanning process is similar to photocopying in that sheets can become skewed going through the scanner, dirt on the glass can result in spots on the images, and the document may need to be rotated 90% to be of most value to users. These functions, which Bielawski and Boyle call "de-skew," "de-speckle," and "clean-up," are performed by the image processing software.

Image capture and processing software needs to offer (Spencer, 1997):

- Detection of double feeds
- Multiple levels of image enhancement
- Document sequence number imprinting control and support
- De-skew capability especially for small documents passed through a high-speed scanner
- Detection of source documents which are blank on the back side (applies to duplex scanners).
- Automatic rotation of scanned images by 90º to 180º.

Optical Character Recognition

Once the digital image is created, it must be indexed in order to be accessible to the user community. There are multiple ways to capture document-specific information that will facilitate future document retrieval. One popular means of capturing content information is to use Optical Character Recognition (OCR) software. Normally, scanned images, which can be thought of as "electronic photocopies," contain imaged representations of text. OCR software searches for embedded text, which exists as patterns in a dot matrix and converts these patterns into text that can be used to index documents for future retrieval.

As Bielawski and Boyle state "Optical Character Recognition is the process of a software application reading the image and interpreting the characters on the page, thereby turning the image into text, making the entire document available for indexing." Once documents are indexed into the EDMS database, captured document information can help a user locate a needed document. The OCR process is very demanding of system resources and can become the bottleneck in the imaging system, operating at a slower speed than powerful scanners.

Profile or metadata information relating to a scanned document also can be manually input to the EDMS database. Metadata information can include items such as drawing description, software version drawing number, scale, revision and title block information. Manual indexing of documents can be a laborious task and efforts must be made to obtain an environment that is conducive to the indexing operation. One such conducive environment is called "heads up indexing." Heads up indexing is accomplished by software that can display both the scanned image and index fields on a screen at the same time. The user can then view the document and enter the metadata information into the database from the same application. This capability enhances the quality of the metadata information entered and improves the manual metadata entry process. To accomplish "heads up indexing," the EDMS and the scan viewing software must be tightly integrated. An example of "heads up indexing" is shown in Figure 4. Note the EDMS metadata fields shown at the bottom of the screen image (Description, Title Block 1 and 2, and Drawing #).

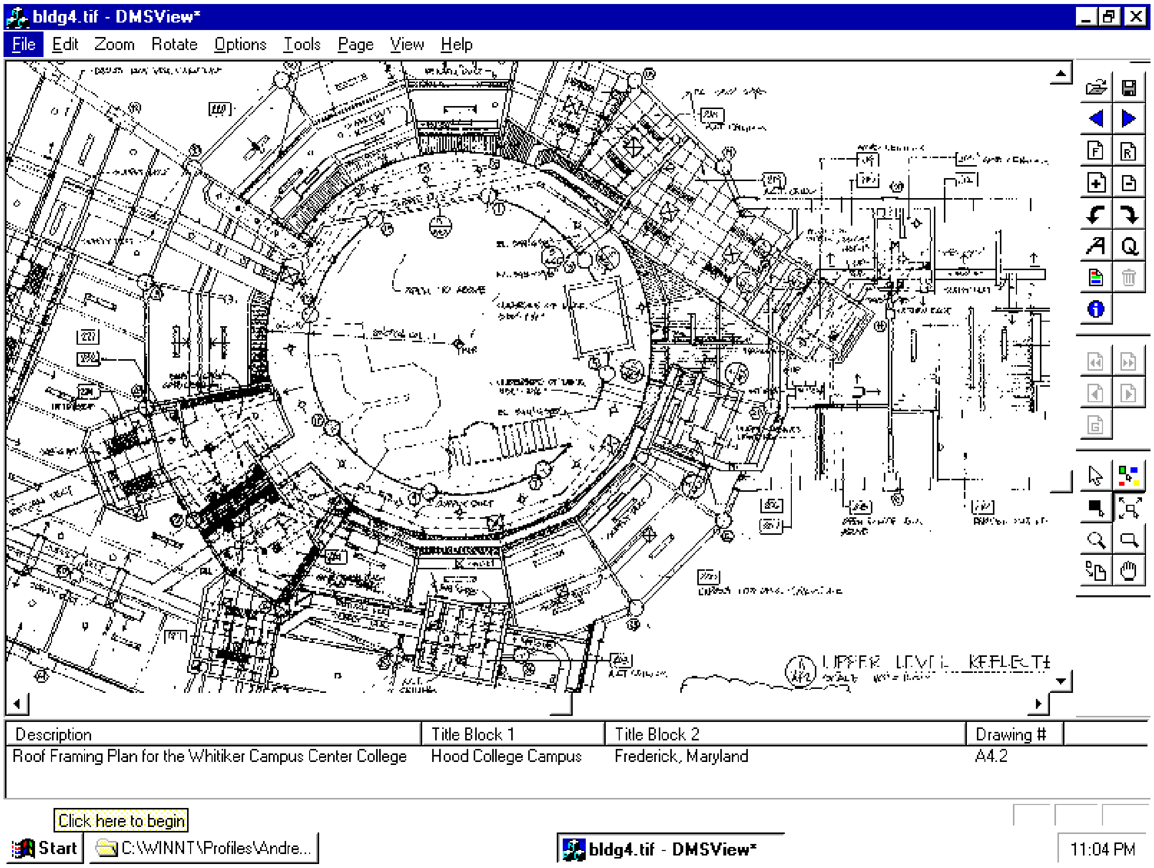

Figure 4. Example of "heads up indexing" of scanned image

Paper Conversion

When considering converting a paper-based engineering document vault into an EDMS-ready digital engineering document vault, questions can be asked to determine the methods available to accomplish the conversion and the viability of the technology in a specific environment.

For engineering and facility management personnel there are a number of methods to get paper-based documents into the automated design and drafting environment that EDMS technology supports. These methods include (Wilson, 1997)

- Manual Redraw – Manual redrawing of a paper document from scratch.
- Digitize – A quicker manual redrawing process using a digitizer tablet.
- Scan – (see Scanning Process)
- Service Bureau – Rather than converting documents with internal resources, some organizations may use external resources to accomplish the conversion process. For example, the Oklahoma City Air Logistics Center has a contract with Boeing Aircraft and ImageMax for the conversion of 238,400 E-3 weapon system aperture cards by January 1999 (Tinker AFB ALC, 1998). For many DoD components, "the Fort Leavenworth Defense

Printing Service Department Office serves as a digitizing/indexing agent, using a high-speed document imaging system" (Smithsonian Institute, 1996). FCAD2 vendors offer scanning subsystem components and access to service bureaus.

Questions which address the need for EDMS technology include:

1. Does your organization have a design review requirement that is heavily paper-based?
2. Is the drawing review process performed by various people?
3. Are documents in the review process generated or distributed externally?
4. Would the review process be faster if documents were viewed simultaneously?
5. Is misfiling or loss of documents a problem for your organization?
6. Do personnel waste time looking for documents or determining availability?
7. Is there a labor intensive process for filing and retrieving documents?
8. Is storage space rapidly dwindling or is new storage space hard to get?
9. Would immediate design availability improve other aspects of your processes?

To develop a ROI (Return on Investment) when considering converting to an EDMS-controlled digital vault from a paper-based engineering database, consider the following questions:

1. On average, how long does it take to file, retrieve, and duplicate a document?
2. How many personnel are dedicated to filing, retrieving, and duplicating paper documents?
3. How much floor space (square feet) is devoted to drawing storage and reproduction?
4. Does your organization produce "paper" copies for distribution?
5. Does your organization produce aperture cards to reduce storage space requirements?
6. Is revenue lost or costs or risks incurred because of lost or missing documents?
7. How much productivity time is lost requesting and waiting for copies of documents?

The EDMS is central to the success of a digitally-enabled paper engineering vault when scanning, indexing, finding and viewing documents. The EDMS is equally important when documents need to be distributed, manipulated, or stored.

Viewing

Users of engineering and facility management documents often do not have access to the native applications used to generate them. To facilitate document access by a wide range of users, viewing software is used. Viewer software allows users to look at documents without having to use the application used to create them (Wilson, Undated). Many document users simply need to see or refer to engineering documents, rather than create or modify them. Viewing software allows such users to access information without having to own and learn costly CADD software packages.

For organizations with a developmental or fully implemented EDMS, “a small investment in a powerful viewing software package offers immediate benefits with little capital outlay and minimal training time (Wilson, Undated).” The primary factors to consider in selecting a viewer software package are (Wilson, Undated):

- Speed
- Simplicity
- Ability to view a full range of file formats

Benefits of Electronic Document Capture

Progressive organizations invest in electronic document capture to realize benefits. Ready access to a repository of electronic documents offers many benefits compared to paper document distribution (Bielawski and Boyle, 1997):

- Lower cost of document distribution
- Easier document maintenance and updates
- Faster access to documents
- Time-consuming linear routing of paper documents avoided
- Views of documents can be customized for a better quality of presentation

Store Activity

Digital information related to the in-process engineering documents include CADD files, scanned images, data used in engineering calculations, word processing documents and spreadsheet documents, e-mail communication, and digital photographs. This information is maintained in digital form on file or data servers.

Metadata

In addition to the actual engineering data, an EDMS system also stores profile or metadata information related to a given document, including title block information, software version, description, revision number and date, document status, keywords, etc. This information is contained in a database system. Because of robust access capabilities and the ability to manage a large number of records, almost all database systems in use today are SQL systems. Therefore, databases are maintained on-line on database servers. Database server functions and file or data server functions sometimes are accomplished via a single hardware server unit.

Storage Components

When engineering documents have reached a released state or are replaced by newer versions, they may be stored off-line on tape or optical media. The information also can be stored on magnetic disk or optical media in a near on-line mode. The level of effort that is required to bring a document back to the in-process state, from a released state, dictates which mode should be used.

The store activity uses the following components. Each of the components encompasses both hardware and software.

- File or data server
- Database server
- FTP server
- Tape Backup (9-track, 8mm, 4mm, etc.)
- Optical Storage (jukebox or individual platter)
- CD-ROM

The role of the EDMS is to store documents in the proper locations on the appropriate devices. Proper, logical storage enables the retrieval of documents for the other activities. Storage and subsequent retrieval need to be accomplished automatically by the EDMS based on logical process steps. For example, when a user makes a request to modify an engineering design document, the EDMS should not require the user to know where the original document resides. Rather, the EDMS should transparently retrieve the document and automatically execute the appropriate application, focusing on the retrieved document. Any related documents, such as MicroStation reference files or AutoCAD XREF files should also be automatically managed with this access request. Conversely, when the user has completed the design session with a particular document, the EDMS should restore the document to its original location without requiring the user to know that location.

Archiving

When documents reach a released state, they may be archived. Archiving entails moving digital information from active hard disk storage to off-line or near on-line storage media, thus freeing up space on active magnetic disk drives for additional in-process documents. The EDMS should be able to manage the archived documents as well as active documents. The EDMS also should control access to the archived documents so that archived documents can not be modified. Typical archive media includes both tape backup and optical storage. Optical jukeboxes, which can store and manage hundreds of gigabytes of digital information as one logical location, have been used for archive storage for many years. Optical jukebox storage access is inherently slower than that of active hard disk storage, but this performance trait historically has been offset by the low cost of storage. However, there have been drastic reductions in the cost of ownership of hard disk drives. Any on-going decision making process involving the economics of optical versus hard disk storage must consider only the most current price information available. Besides its historic cost advantage, optical storage maintains popularity due to the useful life of optical media. Optical storage media can have a life expectancy of 100+ years.

More technically detailed discussion of optical media is included under the sub-heading “EDMS Components – Technical Specifications.”

Tape storage is extremely low cost when compared to other media. Tape storage is popular when immediate or near-immediate access to data is not important to the user. Information is placed on tape sequentially. When a user requests recovery of information from tape, it is conceivable that the operating system would have to read through an entire tape to get to the requested information. This recovery operation could take hours. However, backup tapes are reliable, low cost, and portable. Thus backup tapes are an attractive option in certain environments.

When using either tape backup or optical storage for off-line storage, the EDMS must be responsible for moving the information off-line. The EDMS must also update the metadata associated with each file or document to reflect the archived state. Access to this data should then be controlled by the EDMS so that users may access the data for viewing, but not for modification. Finally, the EDMS should update the related document's metadata to reflect the specific volume that holds the archive. If archived data ever needs to be brought back on-line, the EDMS should be used to find, access, and control this function as well.

Manipulate Activity

Modification of in-process engineering documentation is accomplished via application software packages such as CADD, scan image editing, word processing, e-mail communication, engineering application software packages, and spreadsheets. The processes related to the modification of design information falls within the manipulate activity.

During this activity, workstations are used to modify digital data. Remote users should have access to documents via a dial-up connection, wide area network, or Internet connection. Users responsible for review and redline functions utilize view stations for this activity. Detailed technical recommendations for these components can be found under the sub-heading "EDMS Components – Technical Specifications."

The EDMS should be in full use during all functions related to the manipulate activity, allowing the user to find the document that needs to be manipulated. In doing so the EDMS should also verify that the requesting user has been assigned the proper privileges to allow access to modify the document.

The EDMS also is responsible for the creation of documents. In this activity, certain metadata information such as description, drawing number, user's name etc., should be provided to allow enhanced searching for documents later. The EDMS should ensure that this data has been provided by the user before allowing the function to continue.

The access control of a document is managed by the EDMS during the manipulate activity. When a document relates to an in-process stage in an engineering workflow, users are expected to be able to modify the information as the workflow stage progresses. However, this does not necessarily pertain to all users. The EDMS should be able to differentiate between users and user's functions to control the modification of engineering documents. Therefore, access control is based not only on engineering workflow steps but also on the individual accessing the information.

The EDMS also should be able to manage the passing of a document from one workflow step to another. It also should manage the changing access control related to those workflow steps. Figure 5 shows a typical engineering workflow configuration and the rights associated to a document at the various steps. In this example User 1 is a member of a design team and User 2 is responsible for quality assurance.

Step #	**Document Status**	**User 1 (Design Team Member) Access Control**	**User 2 (Quality Assurance) Access Control**
1	In Revision	Modification	View Only
2	Released for Q/A	View Only	View and Redline
3	In Revision	Modification	View Only
4	Released for Q/A	View Only	View and Redline
5	Released for Construction	View Only	View Only

Figure 5. Workflow and Access Control Example

In this example the document begins at step 1 with a document status of "In Revision." At this time User 1 is expected to be making design modifications to the document. When the document is in the "In Revision" status, User 2, from the quality assurance department, would not be able to modify the document. When User 1 has determined that he/she has completed that step, the document is passed to the next step in the workflow. This step is noted in the above example as "Released for Q/A." At this time User 1 relinquishes the modification rights and User 2 is now able to view and redline the document. In this example, User 2 has determined that User 1 needs to make some changes to the document and changes the status back to "In Revision." This cycle is repeated until User 2 is satisfied with the design. User 2 then changes the status to "Released for Construction." At this step neither user can make further changes to the drawing.

The EDMS should be robust enough to allow for the dynamic changes to a document's status and the resultant access control needs. The EDMS must also take into account the individuals involved at the various steps in the life cycle of a document. Chapter 3 presents additional information on automated workflow.

Distribute Activity

The distribute activity is central to the engineering process for it allows all users to have access to the engineering documents. Those users involved in document related processes need to view paper representations of electronic documents as part of the processes. They can include management personnel acting in a review process, engineering personnel in the field at a construction site, vendors that need access to engineering documents in the form of bid sets, customers or constituents, and sub-contractors involved in the design process.

As the EDMS is central in helping users find, view, and manipulate documents in a controlled and managed environment, it is crucial in the packaging and distribution of the documents.

Distribution of digital engineering documents takes two forms: paper distribution and digital distribution.

Paper Distribution

The EDMS is used in paper distribution through the packaging and distribution of paper documents or document sets. The EDMS is used to create a "print request" and should have the ability to drive associated devices directly. These output devices include small format and large format printers and digital copiers. For many years the technical limitations of these devices prevented the cost-effective and timely production and/or reproduction of engineering outputs. With recent advances in large format print-on-demand systems automated distribution of engineering drawings is possible. Currently available large format output devices (plotters) can produce paper-based representations of engineering drawings at copier speed. In addition, units are available that can provide internal stamping (i.e. Released for Construction), distribution notations on individual pieces, advanced collating and even folding and packaging for manual distribution. While all of these features are valuable in the process, the EDMS must control the features to make them available to users of the system.

The EDMS also is used in the paper distribution to control when documents can be distributed to certain groups based on the current design process step. For example, the EDMS should prevent the production and distribution of documentation to the field before the information is officially released.

Accounting or auditing of the output production should be controlled by the EDMS. The accounting or auditing function should include the generation of transmittal information so that a history of the information distribution is maintained.

Digital Distribution

The EDMS should control and manage electronic or digital distribution, including direct Internet or CD-ROM publishing (such capabilities support creation of Electronic Bid Sets). Digital distribution includes making the electronic information available via the user's LAN/WAN and or the Internet/Intranet. Given that an EDMS allows ready access to documents, controls must ensure the accuracy as well as the availability of the information. Web publishing capabilities illustrate the need to distribute only current, accurate information. There are several ways to distribute electronic documents via the World Wide Web. Available distribution models illustrated on Figure 6 (Boyle, 1997) are:

- Manual
- Publishing
- Access

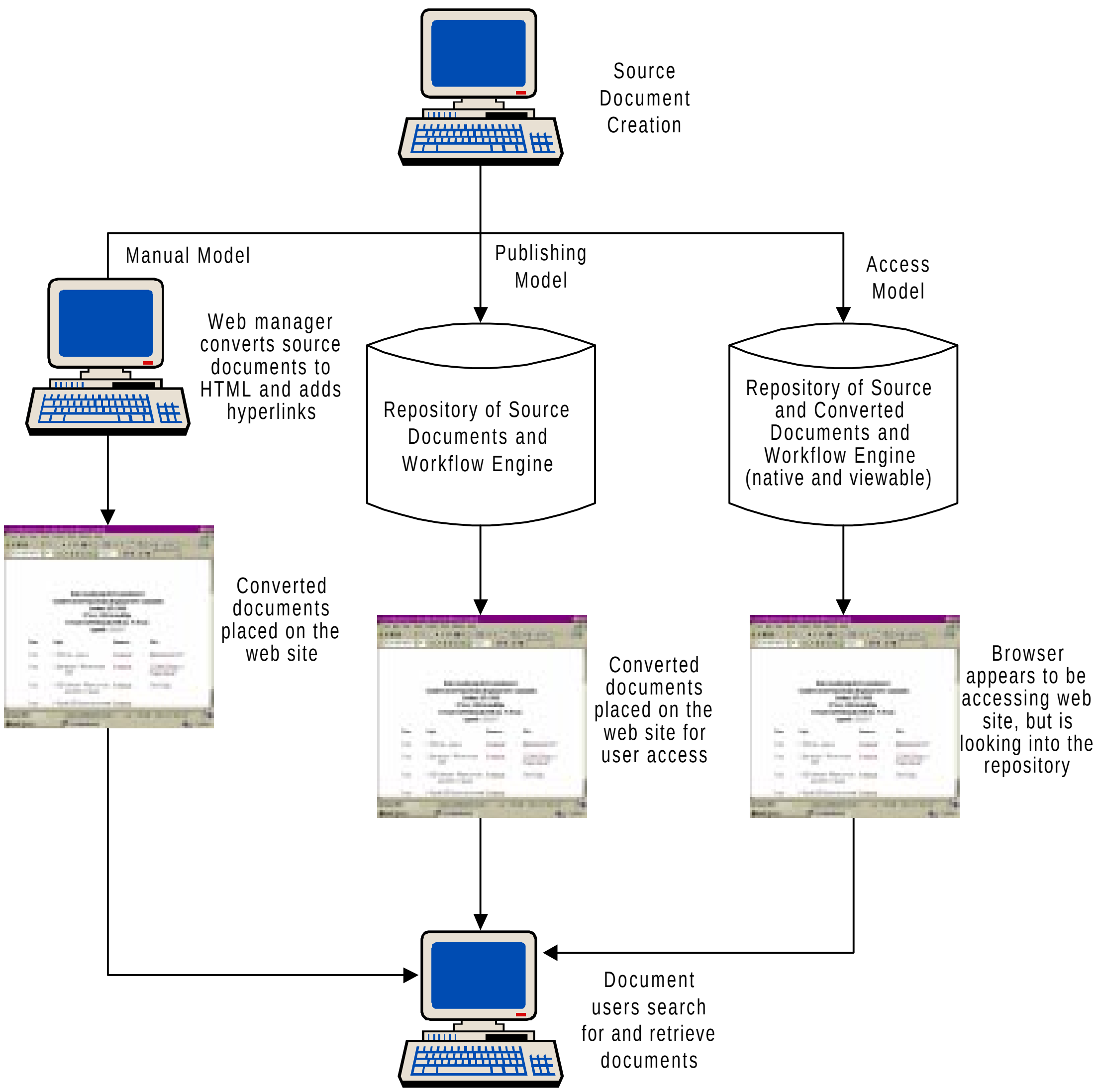

Figure 6. Document Distribution Models (Boyle, 1997)

Each of the three distribution models illustrated in Figure 6 is described below (Boyle, 1997):

Manual Model. Using the Manual Model, source documents are forwarded to a web manager who converts them to a viewable format, adds hyperlinks, and places them on the web site. Document users browse for and retrieve converted documents from the web site. This is a labor-intensive process that has no built in mechanisms to ensure revised source documents get converted and transferred to the web.

Publishing Model. Source documents are stored in a repository and a workflow engine manages the conversion of documents and placement on a web site. Although this model removes the manual conversion and posting process, source documents reside within the repository while converted web-viewable documents reside on the web site (see Figure 7).

Access Model. The access model also uses a document repository and workflow engine. However, the repository is used to store both source (native format) documents and viewable documents. To an end user, the browser appears to access a web site when it actually is looking into the entire repository. Advantages of this model are that search capabilities and security/access controls of the repository are available.

Without an effective EDMS, Web publishing is accomplished without any consideration of the design level or status of the drawing. If a document is in the process of being revised, only those personnel involved in the process should have access to that information. Those involved in resultant construction, or having other uses for information, should not have unlimited access to the document without being aware of its status. The EDMS can and should control this distribution.

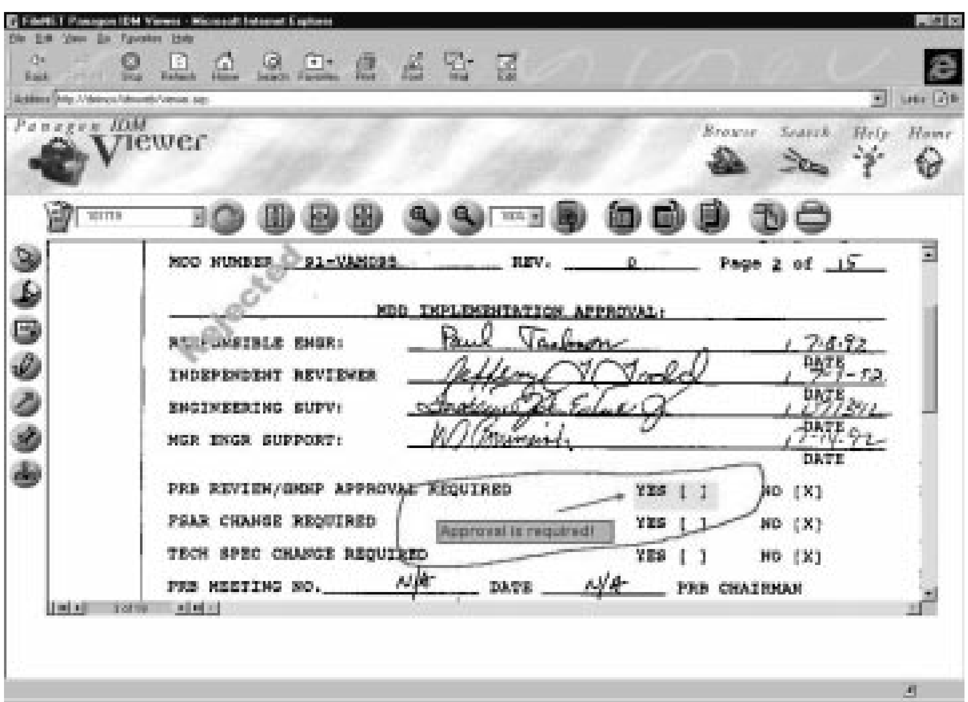

Figure 7. Digital Distribution of an EDMS Document

Also situations exist where released documents would be stored in a different digital format than that of the original document. The reasons for using another format include ease of viewing using standard plug-ins and prevention of released document changes. For example, a designer may use a CADD package such as MicroStation or AutoCAD to create and manipulate the engineering design data. Documents stored in these formats would require complicated and high priced CADD design packages for full document access. However, those persons that only require viewing access to the same information can use an inexpensive and easy to use viewing software package. To view information via a viewing software package the data must be in a standard digital format such as TIFF, CCITT, or PDF. The EDMS should control the creation of the standard digital formats based on the steps in the design process. The EDMS also should control the distribution of this information so that it is available to the viewing user.

Manage Activity

At the heart of the four activities already described lies the manage activity. This is where the EDMS resides. As shown in Figure 1, the EDMS component is responsible for controlling and managing the communications between all the activities.

As the controlling activity, the EDMS is responsible for supporting an organization's workflow. The workflow (routing) for administrative documents, typically prepared by one or a few persons, without extensive reviews and iterations, differs tremendously from workflow for CADD documents. CADD files used for facility management typically are prepared by a number of disciplines and undergo multiple review and revision cycles. For this reason, the remainder of this discussion focuses on CADD document management.

CADD Document Management

CADD document management can be defined as the management and intelligent control of CADD files and their associated references. CADD document management includes documents that are in-process as well as documents that are in a released state. The requirements surrounding the management of CADD or engineering documents are very different from those of administrative documents. Some of the differences include:

- Engineers usually focus on a project and process as opposed to a document.
- The naming of CADD documents needs to be controlled because naming typically ties back to a drawing number. Document numbers have meaning and are pre-assigned in many organizations.
- CADD documents are associated with reference files that must be managed intelligently and automatically.
- The CADD document workflow is extremely dynamic when compared to a typical business workflow associated with business documents.

- Vendor and/or customer communications require native CADD data to be transferred for common editing.

- Design data can be object based. Native CADD encourages object management even in the most basic CADD usage (i.e. MicroStation cells and Autodesk blocks). More intelligent data may be assigned to individual elements.

- Management of CADD documents has been performed manually for many years. However, this manual control did not prevent multiple access to a representation of the same document. Nor did it recognize drawing process states that could affect access control of the information.

- The output of design data in many cases is not portrayed as WYSIWYG. Rather it is displayed as raw data or model data. Only when the design data is passed through the plotting, printing, or publishing sub-system does the resultant image (i.e. plot) emerge. Typical business documents are created and remain in a WYSIWYG environment.

Benefits of CADD File Management

- Provide CADD file consistency through standardized drawing creation methods

  The importance of standardization in document creation is especially important when CADD documents are involved. In many cases a CADD document is part of a larger set of documents or drawings. All of the documents should be subject to the same constraints. These constraints include drawing border selection, population of the drawing title block data, drawing units (i.e. feet, inches, etc.), file and document naming, base reference information, and file location.

- Provide a means to measure and analyze revision history

  The EDMS system should allow the user or document owner the opportunity to study the access history of a drawing. This capability allows the user to gain knowledge about the workflow that took place during the design and use life of the drawing. Information regarding the cumulative amount of time spent in the creation and modification of the drawing is also very useful in forecasting and budgeting future efforts.

  The collection of accounting or audit information should happen automatically without any intervention from the user. The information that is to be collected should be definable by the system manager. This definition should include the ability to include any document related metadata information such as keywords, phases, status, description, and title block information for each activity accounted for. Accounting or audit information is especially useful in studying a document's workflow.

  The historical data should be captured in standard SQL data tables or as delimited ASCII information for import to other systems.

- Eliminate "dead time" in searching for a drawing and verifying that you have the latest revision

  One of the main benefits of an EDMS is the help that it gives the user in locating a drawing when the user wishes to view or edit the document. The EDMS provides an efficient way of listing drawings that fit a user defined criteria such as project and discipline. In addition the EDMS allows the user to perform more extensive searches when required. Examples of more extensive searches include searching based on keywords, revisions, and even words contained in a description or title block line.

  The EDMS also ensures that the user is given modification access to only the latest revision or version of a drawing. Version control capabilities guarantee that the user's modifications are made based on the latest drawing content and that recreation of design data will not be required.

- Eliminate "dead time" spent on recreating design information because finding a lost document is uncertain

  When an engineer, designer, or draftsperson has difficulty finding an electronic document, the individual has no way to determine how long it will take to find the document or even if the document exists any more. In such situations it is not uncommon for drawings to be recreated rather than continue to search for them. An effective EDMS can completely eliminate this wasteful use of time.

- Prevent multiple simultaneous revisions to the same document

  Sometimes keeping several versions of an electronic drawing is desirable in order to maintain design information at various stages throughout the design process. Therefore, a user could access, for modification, a version of an electronic drawing that may not be the latest version or revision. If an old version was retrieved, modifications will be lost and the time spent making the changes will be lost as well. Also the information delineating the design effort at a specific version or revision level has been compromised. An effective EDMS will prevent such problems.

Basic Requirements for CADD File Management

- Provide an easy to use user interface that matches the user's process environment

  The user's interface should be a GUI (graphical user interface) and should be easy to use. When an EDMS system is installed and started up, gaining the user's acceptance of the system is important. The business impact of using the EDMS system can be profound but can be a failure if the cultural impact of the system's use is not taken into account. The user interface should match as closely as possible the manual access by the user to minimize the cultural impact that the system will have. The use of the EDMS should be viewed by the user as a tool and not as another application that they need to learn, execute, and maintain. Therefore the EDMS should be (or appear to the user to be) integrated into the CADD applications. For example, when the user is running either MicroStation or AutoCAD, the EDMS "open document" function should appear on the screen when the user invokes the "open file" option on the standard menu.

- Enable two-way linkage between the database and the CADD drawing (intelligent data linkage)

  As users begin to depend on the EDMS search capability, the metadata and the related title block information must match.

  The EDMS should have the capability to place standard CADD text onto the drawing to match the metadata that is placed into the EDMS document database record. Metadata fields such as drawing number, revision, revision date, title, drawn by, etc., can typically be found within the title block of a drawing. The user must be able to search the database using these items. Therefore the data on the drawing itself match the data in the database record for that drawing. The software version that created the electronic file is critical for long term data storage and should be a metadata requirement for each file. CADD, especially, will require software tracking since successful plotting of a CADD file may require the right combination and version of CADD, raster and plotting software.

  Data linkage must be bi-directional. For example, if the metadata information is modified in the database, the EDMS should automatically modify the core drawing file. Also, if the user changes the text in the CADD drawing file using the text editing tools provided by the CADD software, the metadata information related to the drawing file should update automatically.

- Provide features to facilitate client packaging issues

  The EDMS should have the capability of packaging CADD and related documents for delivery to outside concerns such as vendors, contractors and customers. The user should be able to:
  - Select the documents to be transferred
  - Define the person and organization that is to receive the information
  - Determine if the documents to be transferred are to be released for modification or for view-only. If the documents are transferred as modifiable files (i.e. the recipient is expected to modify the design information), the EDMS should then set the status in the database so that local users may not modify the files.
  - Combine all the files into one document using standard compression techniques (such as Pkware's PKZIP product)
  - Create a transmittal showing the file name, drawing number, and a description as a minimum
  - Define a note to be included in the transmittal in order to pass additional information to the recipient regarding the package
  - Provide an audit trail to delineate what files have been transferred, when

they were transferred, and what location and user they were transferred to.
- Control access to EDMS functions on a user or user group basis.

The system administrator should be able to define attributes about an individual user or group of users that are related to the user's access to documents and the operations that can be performed on documents. In some cases EDMS systems can dynamically modify the GUI used for document access based on the user's rights or permissions as defined by the system administrator.

- Provide the capability for the system administrator to define document status parameters and associated access control

  The EDMS should have the ability to define a document status that controls access to a particular document. For example, a document that has a status of "Released for Construction" should modify the access control of that document, so that modification to the design data is not permitted. However, in this case, viewing (not modifying) the document should be allowed. System administrators must be able to define their own status labels to match the workflow steps that are employed in their organization.

- Interface with existing plotting sub-systems and capabilities

  Plotting is a very important function of CADD operations. This function provides the ability to create a paper copy of the digital information. The EDMS system should be able to integrate into whatever plotting system is selected by the user. The EDMS also must maintain an accounting of when plots have been created and who requested the production of the plotted output.

- Enable global database changes (i.e. revision, status, title block, etc.)

  In many cases modification of a metadata field or fields based on process changes is necessary. For example, when a project is released for construction, the user should be able to request that the status of all of the documents related to the project be changed to "Released for Construction." The EDMS should provide this feature which should allow the user to select the set of documents that would be affected and the content of the metadata field(s) to be modified. The metadata field(s) for all of the selected documents should be updated with a single operation.

- Facilitate standard drawing creation

  An EDMS must have the capability to define and enforce a standardized drawing creation function that is based on customer defined criteria. The EDMS should either have this capability built into the product or allow programmatic input to be defined. This feature should be definable on a project basis or environment basis.

- Allow for electronic sets or folders

The EDMS should have the capability to group files together into logical grouping called electronic sets or folders. The grouping should act as one unit for all operations in the EDMS including check-in, check-out, move, copy, etc. Many engineering design software products that operate with the CADD software create multiple files such as configuration files and database files. In these cases all related files are required for the correct operation of the engineering design software products. The EDMS should have the ability to manage all the related files and should automatically populate the EDMS database when related files are created.

- Provide control of related non-CADD documents (i.e. client specifications, project meeting notes, vendor specifications, etc.)

  The EDMS should automatically manage non-CADD format documents that are related to a specific CADD document and to the engineering process. Examples of these documents include client specifications, design specifications, project meeting notes, vendor specifications, etc.

- Provide security at the drawing level as well as the user level while respecting network security specifications

  The EDMS should allow both file level security as provided by the operating system in use and process level security as defined by project managers from within the EDMS. The EDMS should not dictate the extent of each type of security, but permit the system manager define the levels of protection.

  The EDMS security at the drawing level is used to protect the document content based on the process or workflow step at a given point. For example, if a document has been released for quality assurance, only those individuals in the quality assurance roles should be able to redline the drawing. At the same time, all other users should be allowed view only access to the document.

  Security at the user level enables the system administrator to define what EDMS operations individual users or groups of users may perform. For example, a novice user may be limited to viewing CADD drawings. In this case the EDMS should prevent the user from modifying the CADD document.

- Control the create, rename, copy, copy and replace, delete, and move document functions

  To control document operations on an individual operation basis, user level security should be configurable within the EDMS.

- Enable revision control

  The EDMS should have the ability to create and manage revisions or new versions of a CADD document. The EDMS system manager should be able to define how revisions or versions are named in order to set them apart from the latest revision or older revisions.

The document status should be definable based on revision changes and storage. Since status control could affect access to a document, it can be used as a way to prevent the modification of archived revisions or versions of a CADD document.

Revision or version control definition should be done on a document format basis. Some document types require revision control and monitoring and other document types require no revision control. When revisions can be controlled on a document format basis, those formats that require little or no revision or version control do not suffer from the system overhead required by document format types that require revision or version control.

- Automatically control and manage all reference file information

  Many popular CADD systems allow the use of reference or overlay information. In AutoCAD these files are called XREF files and in MicroStation these files are called reference files. The EDMS should recognize when reference files are associated with a CADD document and should satisfy the need to display the reference files without any user intervention. The EDMS also should recognize associated CADD support files including pen tables, color tables, and font resource files.

EDMS Components – Technical Specifications

Network Hardware

Network hardware is the backbone of any EDMS system. It carries all digital data from client node to server, server to client, as well as from client to client. Hardware components of any network include a network interface card (NIC) within each node, various cabling, and at least one node acting as either a dedicated server or both client and server. In larger networks, repeaters and hub devices can be used to “boost” and distribute the network data transmission. The collection of equipment functioning as a single communication group is traditionally referred to as a Local Area Network or **LAN** (Figure 8). LANs are designed to enable users to create, view, and modify centrally located documents that are stored on the “server” platform. Multiple users can simultaneously access these documents. Network operating system, EDMS software, and a database system are installed onto the server hardware to provide and manage these capabilities as well as other server functions including faxing, printer sharing, Internet publishing, and database storage.

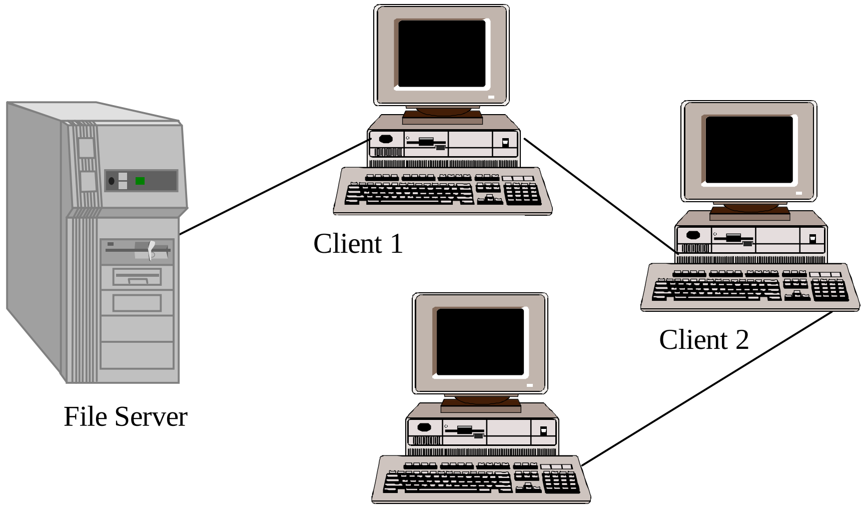

Figure 8. Basic PC Local Area Network Configuration

Network hardware devices adhere to certain specifications used to classify their compatibility and performance. The most common hardware is known as Ethernet. Ethernets can transfer data at a rate of up to 10 megabits per second (mbit/sec). Ethernet networks operate in a “highway” environment in that data transmitted to or from a client competes for the usage of a common medium. Therefore, when Ethernet communication traffic is heavy, the highway can become congested and clients may have to wait to use it. Network congestion causes a loss in performance. Frequent congestion occurrences over a small amount of time can become a serious bottleneck. With the addition of more clients transferring data on a LAN simultaneously, the network’s performance will further diminish. Performance loss is exacerbated in the engineering process in that engineering documents (i.e. CADD files) tend to be much larger than documents related to the business process in a properly designed system. EDMS software can manage the engineering process and workflow to minimize this traffic.

Within the last few years, two new Ethernet solutions have been introduced to improve performance. The first solution is known as Ethernet Switching and is shown in Figure 9. Ethernet switches are hub-type devices that dramatically reduce the occurrences of data competing for the same limited bandwidth. Ethernet switches help to maintain the performance of the network as data traffic increases. A high speed “switching engine” chip is used to intelligently route network data from source node to destination node without broadcasting (repeating) it to the other nodes within the LAN. Basically, each node is given its own “dedicated lane” within the highway on which it flows. The second improvement to Ethernet has been the introduction of Fast Ethernet. Fast Ethernet works the same as Ethernet except that it transfers data at 100 mbit/sec, 10 times faster than Ethernet (10 mbit/sec). It is important to note that Ethernet and Fast Ethernet nodes cannot communicate with each other directly. An Ethernet/Fast-Ethernet Switch is required to enable the two technologies to coexist. Ethernet networks rely on either fiber-optic, twisted pair, or coaxial cable to function. Fast-Ethernet networks require either Category 5 (CAT5) twisted-pair or Fast Ethernet fiber-optic cable and transceivers. Additionally, Fast-Ethernet NICs are able to perform at regular Ethernet speed (10 mbit/sec), but Ethernet NICs are not able to function at Fast-Ethernet (100 mbit/sec) speed.

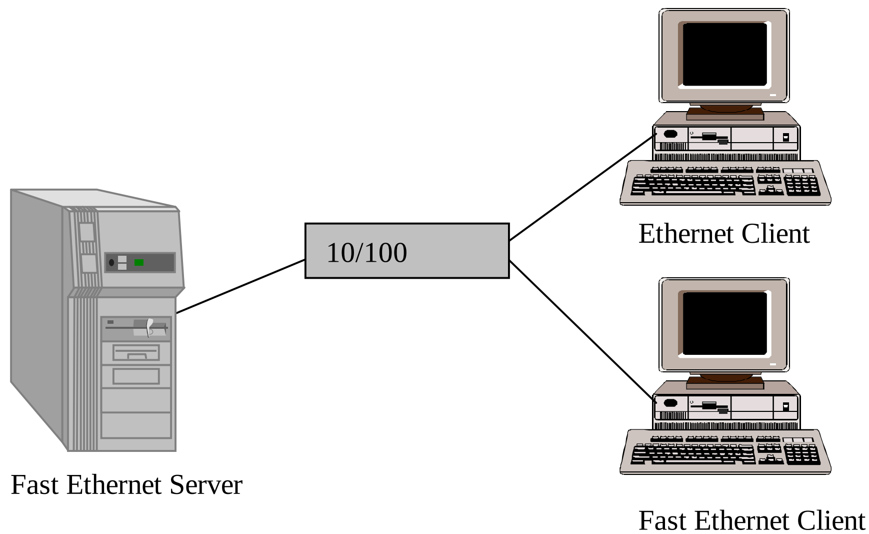

Figure 9. Modern Ethernet/Fast-Ethernet LAN

Server Node

The server node can assume many roles in the implementation of an EDMS system. The LAN server ideally is located in a central location that users access for any of the following resources:

- databases
- documents
- applications
- printing
- faxing
- web publishing
- central data backup
- data archival

The role of the server is not always as a dedicated one. It is possible with commonly available operating systems and hardware, to create a small network that will use PC hardware as both client and server to the network. Typically, networks rely on either a single PC or many PCs to take on the dedicated role of server. Within the realm of EDMS, the mission critical nature and the vast amount of documents being managed require a server to be secure, fault tolerant, and highly available. These features are common and easily acquired from today's systems vendors.

Currently, three Network Operating Systems are most commonly used within EDMS configurations: Novell Netware, various versions of UNIX, and Microsoft Windows NT Server. Each of these systems can provide:

- File and printer sharing
- Database hosting
- Internet publishing
- Primary and secondary storage
- Fault tolerance
- RAID level data guarding

The hardware requirements associated with such features vary with each operating system product and the hardware to be integrated.

Currently, Microsoft's Windows NT Server is capturing most of the market for new server installations while most hardware manufacturers and software developers are selecting it as their primary platform for new product development. The following is a list of some of the reasons why this product has gained such popularity:

- Low cost of ownership including purchase, implementation, and training
- Bundled Internet tools
- Uses common Graphical User Interface (GUI)
- Administrative Wizards for novice administrators
- Strong network connectivity with many other networks
- Flexible client licensing schemes
- Large choice of available server applications: SQL, e-mail, Internet, document management, file sharing
- Large choice of compatible hardware
- Bundled migration tools
- Fault tolerance support

Due to its popularity and emerging prominence, Microsoft Windows NT will be used as the Network Operating System within the various illustrations and explanations throughout this section on EDMS components.

System Requirements

EDMS servers utilizing Windows NT Server V4.0 as the Network Operating System demand a large amount of storage space as well as RAM memory. Windows NT Server, like other multi-tasking operating systems, performs at its best when it does not have to write memory pages out to disk when memory utilization is maximized. Windows NT Server also takes advantage of extra available memory to map contents of frequently accessed files directly to memory, thereby boosting the reading/writing performance of those files.

Processor type and speed must be considered as well. Windows NT Server supports Symmetric Multi-processing (SMP) and therefore can utilize multiple processors to further boost its processing performance. While performing typical server tasks such as file and print operations, SMP would most likely not be a significant advantage. However, when EDMS related server applications such as SQL database, network faxing, and/or e-mail are to be integrated, then SMP could be a benefit since many more processing operations would occur simultaneously.

Important note: Not all processors are available in SMP configurations. Currently, the Intel Pentium II processor in multi-processor configurations is only available from a few systems vendors. The Intel Pentium Pro is currently considered the workhorse processor in SMP configurations. All major brand name server vendors currently support SMP hardware with Windows NT Server V4.0 as a preinstalled Network Operating System.

The following guidelines (Figure 10) can be used to determine the minimum server requirements to support EDMS.

	Windows NT Server V4.0
MINIMUM SINGLE PROCESSOR	Pentium II 233 mhz
SMP processors	Pentium PRO or Pentium II
Minimum memory	128 MB
Minimum hard disk space for operating system	1.0 GB
File system	NTFS
Network interface	Fast Ethernet or Gigabit Ethernet
Primary storage type	RAID 5 HW based hot pluggable

Figure 10. EDMS Server Base System Requirements

A suggested formula for determining storage requirements is:

TOTAL PRIMARY STORAGE =
(total size of server apps + NOS (**N**etwork **O**perating **S**ystem) size + max print spooler size + total page file size + existing documents size + existing database size + future database size for future added docs + anticipated future added documents size) *1.30

NOTE: 1.30 Multiplier is used to maintain the primary storage at a maximum 70% full rate to reduce disk fragmentation.

Input/Output System

The size of an electronic document to be stored is affected by certain variables. These variables include the physical size of a document, the resolution of the scanned image, and the storage of the image as color, grayscale or black/white. Therefore, documents in an EDMS have no “typical” file size.

The core function of any server is the transfer of various types of data to and from various components within the server as well as externally to clients throughout the LAN, WAN, or Intranet and Internet. Application and document access as well as printing and database manipulation all demand substantial resources from the input/output (I/O) system. The selection of a quality high performing and expandable I/O system components is one of the most important decisions that must be made, especially when considering the large amounts of data that are involved with the engineering process. Often, the decision to select the fastest CPU processor speed is mistakenly made and consideration of the I/O system is overlooked. File and print data account for the majority of the system bus utilization within a standard server. I/O system components include memory, hard disks, disk controllers, system bus, and network interface cards.

The various storage systems within an EDMS server all lie within the I/O system and all compete for its resources. These storage systems are as follows: Primary, Secondary, Magneto-Optical (MO) Storage, and Tape/Sequential. Each of these systems has pros and cons regarding their usage, performance, and associated costs. Figure 11 depicts the relative costs and performance as well as the unique characteristics of each system. In addition, each defined storage system can be used to compliment the other as far as redundancy and protection of data, which is of extreme importance when maintaining any EDMS.

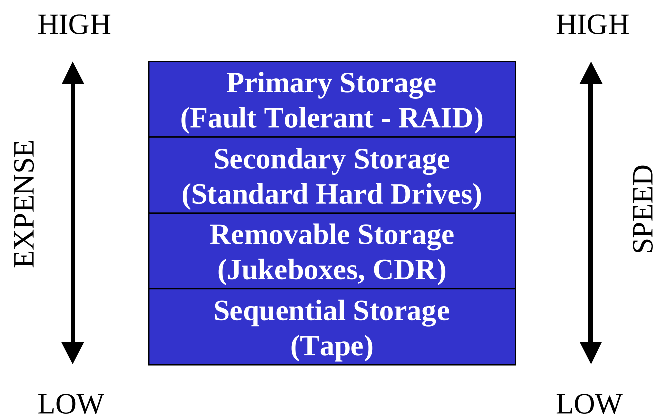

Figure 11. Storage Management Hierarchy

Primary Storage and RAID

RAID (Redundant Array of Inexpensive Disks) sets are used in many EDMS servers and give fault tolerant high availability to data servers. Of the five levels of RAID, RAID level 5 is the most popular and recommended for EDMS server's document storage. RAID 5 enables a set of disks to act as a single logical partition and uses parity information (called "parity strips") to protect data should a single disk failure occur. RAID 5 dedicates the equivalent of one disk for storing the parity strips but distributes the parity strips across all the drives in the group. The data and parity information are arranged on the disk array so that they are always on different disks. A minimum of three disks is required to create a RAID 5 set. Both software and hardware versions of RAID 5 support are available. To maintain the array of disks without impacting resources from the central I/O system the recommended approach is to use a hardware solution since there is dedicated processor and memory included on RAID 5 disk controllers.

Secondary Storage

A secondary storage system utilizes standard hard disk technology which does not provide redundancy should a component fail. Current hard disk technology provides massive storage capability and performance at a very nominal cost. An EDMS server could contain both RAID 5 and secondary storage drive sets. Non-critical documents could be placed onto secondary storage, and performance on par with primary storage could be maintained at a cost significantly less than placing non-critical documents onto primary storage.

Removable MO Storage

Removable storage devices function similar to a hard disk, yet the media is removable and can have a life expectancy of 100+ years. The most popular attribute of these storage systems is their ability to separate the media from the storage unit and place it in off-line storage. This function can be automated if an optical disk jukebox device is integrated into an EDMS server. These devices have robot-type mechanical components that move media in and out of the drive mechanism and place the ejected media onto the unit's internal media rack or into a "mail slot" for removal. Currently, units can store up to 600 gigabytes of data or approximately 155,000,000 pages of text or 2,500,000 CADD drawings. Access time for these devices is much slower than for hard disks because of the technology used and also the fact that documents requested may be located on a stored cartridge that must be loaded into the drive mechanism prior to document retrieval.

Network operating systems do not supply software drivers to manage jukeboxes. Therefore, a software solution is required when integrating one of these units. Jukebox management software falls into a category known as HSM (Hierarchical Storage Management) which can automate the process of moving aged or inactive documents onto these devices to free primary and secondary storage for active data. With these systems, EDMS users do not need know the physical location of their documents. Retrieval is performed automatically when the documents are requested through the server operating system. HSM systems deliver the management function that migrates aged and inactive documents to near online storage (MO jukeboxes) so that the primary storage will run efficiently and not be congested with documents that are not frequently accessed. HSM is a software solution that typically can access all types of storage hardware and therefore can be utilized within very diverse hardware configurations.

COLD Technology

Computer Output to Laser Disk (COLD) technology is confused at times with the imaging world. COLD is used in organizations that require computer systems to generate large volumes of data and reports. COLD is typically not used for in-process engineering documents but rather letter/legal size business documents. The traditional solution has been COM (Computer Output to Microfilm) systems, which record general computer documents onto microfilm.

With COLD technology, computer reports that would otherwise be printed on or sent to a microfilm recorder are instead transferred directly to an optical disk. Most COLD systems use WORM (Write Once Read Many) disks since WORM provides the large storage capacity and permanence that archived records normally require. Because COLD and image technology both use optical disk as the storage medium, sometimes the two are confused. The key distinction is that document imaging stores replicas of paper based documents that have been scanned into the system, while COLD stores raw computer data put directly into a storage application and then displays this data in a variety of report formats. COLD also automatically indexes incoming data. COLD systems are well suited to organizations that create, maintain, research, and store high volumes of statements, invoices, checks, account status documents, or other types of computer generated financial records

Benefits of COLD

- Price/Performance - low cost, high storage capacity
- Fast Access - a significant improvement over paper
- Multiple Field Indexing - Well designed software lets users locate data by any indexed key.
- Text Search - Users can search for text strings automatically, allowing retrieval of specific on-line data.
- Archivability - Optical disks have a long life well beyond 30 years.
- Legality - Rulings have determined that optically stored records are true copies of the original data.
- Return on Investment - Most COLD systems have a pay-back period of less than six months.

Worm Technology

WORM technology (Write Once Read Many) is hardware and/or media that assures that data can be written to the media only once and cannot be removed. CD-ROM, CD-R, and WORM MO media are examples of this type of technology. The rigid capability of only one writing pass to the disk track is beneficial to corporations that require archival data to be safe from possible overwriting. This technology, just as other mediums, is not tamper-proof or protected from physical misuse and damage. However, WORM media is considered to have an archival life of at least 30 years.

Sequential Storage

Sequential storage systems are low performance devices used for data that is being purged from the system due to obsolescence. Since data is stored on these devices sequentially, they are considered to be off-line from access by users. Usually, an administrator would be required to manually recover the data should it become necessary to again activate its use. Typically the media from a sequential storage device (tape or cartridge) is stored off site. Off site storage protects the data should a catastrophe such as fire occur. These devices may require additional software to make them function with the network operating system.

Currently various sequential systems are available. Very high capacity devices use mechanisms similar to MO Jukebox devices that maintain a set of tapes as single logical tape. Such devices provide automatic tape insertion during backup and restore operations without human intervention. Figure 12 shows capacity and throughput specifications of four commonly used tape storage systems.

	TRAVAN	**4mm DAT**	**8mm Exabyte**	**DLT tape**
Max capacity/tape	20 GB	96 GB	3,200 GB	3,400 GB
MAX THROUGHPUT	600 Kbyte/Sec	2,200 Kbyte/Sec	25,000 Kbyte/Sec	20,000 Kbyte/Sec

Figure 12. Popular Tape Backup Device Specifications